

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: October 19, 2002, 06:53:36 ; Search time 25 Seconds  
(without alignments) 109.342 Million cell updates/sec

Title: US-09-807-459-2  
Perfect score: 2359

Sequence: 1 MAPSDSVDVTKILLAASES.....DPSKALIRKVSTEADNLLEK 458

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

## Searched:

10524 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : SwissProt\_40.0\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	127	5.4	607	1 YMHL_YEAST	YMHL_YEAST / STANDARD; PRT; 607 AA.
2	122	5.2	1829	1 FPRC_NEIMC	004632 saccharomyces_neimic
3	121	5.1	566	1 HEMA_IAUSS	P55127 neisseria_m
4	120.5	5.1	566	1 HEMA_IAZU	P03453 influenza_a
5	117	5.0	1038	1 C1NB YEAST	P26141 influenza_a
6	117	5.0	1829	1 FPRC_NEIMB	P27895 saccharomyces_neimic
7	116	4.9	457	1 V162_IUREPA	Q9JYV5 neisseria_m
8	115	4.9	1875	1 MLP1_YEAST	Q9PQY2 uracilamino_acid
9	113.5	4.8	2116	1 MYS2_DICDI	P08799 dictyosteliidae
10	113.5	4.8	2459	1 MAPB_RAT	P15205 rattus_norvegicus
11	112	4.7	566	1 HEMA_IALEN	P18876 influenza_a
12	112	4.7	1957	1 YD86_SCHPO	Q10411 schizosaccharomyces_pombe
13	111.5	4.7	832	1 ALP6_SCHPO	Q9usq2 schizosaccharomyces_pombe
14	110.5	4.7	886	1 ORCL1_KUDLA	P54798 klyveromyces
15	110.5	4.7	6359	1 DACC_BACLT	068088 bacitracin
16	110	4.7	566	1 HEMA_IARPE	P03452 influenza_a
17	110	4.7	577	1 TIPM1_SCHPO	Q09684 schizosaccharomyces_pombe
18	110	4.7	1 Y338_MYCGE	P47580 mycoplasma_giganteum	
19	108	4.6	566	1 HEMA_IARKE	P18815 influenza_a
20	107.5	4.6	462	1 YB47_METJA	Q58547 methanococcus
21	107	4.5	550	1 HEMA_IAZH3	P11334 influenza_a
22	106.5	4.5	4036	1 RRLPL_DUGB	Q66431 du邦 virus
23	106	4.5	550	1 HEMA_IADH6	P12587 influenza_a
24	106	4.5	550	1 HEMA_IADH7	P12588 influenza_a
25	106	4.5	624	1 YM08_YEAST	P54074 saccharomyces_cerevisiae
26	105.5	4.5	550	1 HEMA_IADHK	P43257 influenza_a
27	105.5	4.5	566	1 HEMA_IATEN	P03440 influenza_a
28	105.5	4.5	1030	1 Y075_MCPN	P75556 mycoplasma_cryptogene
29	105.5	4.5	2670	1 YAO5_SCHPO	Q10105 schizosaccharomyces_pombe
30	105	4.5	808	1 Y066_NPVC	P41457 autographa
31	4.5	857	1 NEM_CHICK	P16053 gallus_gallus	
32	4.4	778	1 YF05_METJA	Q58900 methanococcus	
33	4.4	550	1 HEMA_IADH1	P12582 influenza_a	

## ALIGNMENTS

P12583 influenza\_a  
P12584 influenza\_a  
P14873 mus musculus  
P26135 influenza\_a  
P47808 acanthamoeba  
P08096 schizosaccharomyces\_pombe  
P06075 mus musculus  
P36112 saccharomyces\_cerevisiae  
P53705 candida\_albicans  
P03449 influenza\_a  
P32497 saccharomyces\_cerevisiae  
Q00402 saccharomyces\_cerevisiae

P12583 influenza\_a  
P12584 influenza\_a  
P14873 mus musculus  
P26135 influenza\_a  
P47808 acanthamoeba  
P08096 schizosaccharomyces\_pombe  
P06075 mus musculus  
P36112 saccharomyces\_cerevisiae  
P53705 candida\_albicans  
P03449 influenza\_a  
P32497 saccharomyces\_cerevisiae  
Q00402 saccharomyces\_cerevisiae





Query Match 5.0%; Score 117; DB 1; Length 1038;

Result 5

CIN8\_YEAST STANDARD PRT; 1038 AA.

AC P27895; DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-FEB-1995 (Rel. 31, last sequence update)  
DT 15-JUL-1998 (Rel. 36, last annotation update)

DE Kinesin-like protein CIN8.

CIN8 OR KSL2 OR YEL061C.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI\_TAXID=4932;

RN [1]

RP SEQUENCE FROM N.A.

STRAIN=S28C / AB372;

RA Hoyt M.A., He L., Loo K.M., Saunders W.S.;  
RT "Two Saccharomyces cerevisiae kinesin-related gene products required  
for mitotic spindle assembly.";  
RL J. Cell Biol. 118:109-120(1992).

RL [2]

RP SEQUENCE FROM N.A.

STRAIN=S28C / AB372;

RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,  
Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,  
Chung E., Duncan M., Guzman E., Hartzel G., Hunnicke Smith S.,  
Komp C., Lasikari D., Lew H., Lin D.,  
Mosedale D., Nakahara K., Namath A., Norgren R., Oeffner P., Oh C.,  
Petei P.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,  
Taylor P., Wei Y., Yelton M., Boström D., Davis R.W.;

RA Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 1021-1038 FROM N.A.

RC STRAIN=S28C;

RA MEDLINE=9517238; PubMed=7867803;

RA Rousset G., Simon M., Riposte P., Buhler J.M.;

RT "A second nitrogen permease regulator in *Saccharomyces cerevisiae*.";  
RL FEBS Lett. 359:215-219(1995).

RL [4]

RP CHARACTERIZATION.

RC STRAIN=S28C;

RA MEDLINE=92354062; PubMed=16436659;

RA Saunders W.S., Hoyt M.A.;

RT "Kinesin-related proteins required for structural integrity of the  
mitotic spindle.";  
RL Cell 70:451-458(1992).

CC -1- FUNCTION: REQUIRED FOR ASSEMBLY OF THE MITOTIC SPINDLE. INTERACT  
WITH SPINDLE MICROTUBULES TO PRODUCE AN OUTWARDLY DIRECTED  
FORCE ACTING UPON THE POLES. FOLLOWING SPINDLE ASSEMBLY, CTN8 AND  
KIPL APPARENTLY ACT TO OPPOSE A FORCE THAT DRAWS SEPARATED POLES  
BACK TOGETHER. THIS FORCE SEEMS TO BE MEDIATE BY KAR3.

CC -1- SUBCELLULAR LOCATION: SPINDLE MICROTUBULES THAT LIE BETWEEN THE  
POLES.

CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BIMC  
SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
or send an email to license@isb-sib.ch).

CC -----

DR Z11859; CAAT7885.1; --

DR EMBL: M90522; AAA44961; --

DR EMBL: U18795; AB05026.1; --

DR EMBL: X9105; CAA55722.1; --

DR PIR: B42641; B42641.

DR HSP; P1719; 3KAR

DR SGD; S0000787; CIN8.

DR InterPro; IPR001752; kinesin.

DR Pfam; PF00225; kinesin\_1.

DR PRINTS; PR00380; KINESINHEAVY.

DR SMART; SM00129; KIS\_C; 1.

DR PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; 1.

DR PROSITE; PS50067; KINESIN\_MOTOR\_DOMAIN2; 1.

DR Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;  
KW Mitosis; Cell cycle.

FT DOMAIN 72 553 KINESIN-MOTOR (BY SIMILARITY).

FT DOMAIN 554 657 COILED COIL (POTENTIAL).

FT DOMAIN 904 942 COILED COIL (POTENTIAL).

FT DOMAIN 166 173 ATP (BY SIMILARITY).

FT NP\_BIND 254 D -> A (IN REF. 1).

FT CONFLICT 831 831 Q -> H (IN REF. 1).

FT CONFLICT 1038 AA; 117999 MW; 3ALFDY003EF89FBC CRC64;

RP STRAINS-S28C;



Best Local Similarity 20.2%; Pred. No. 4.9;  
Matches 81; Conservative 62; Mismatches 150; Indels 108; Gaps 19;

Qy 94 LPPDQVQLEAFNLFKESDANPANSSTEKFRRMFRRGKHSYFHDLYVNLLERKNTRDADA 153  
Db 35 LGDRYADDAAAIVGDKANLNLN---LNM---KKG-----VENLMDDTGK--K 75

Qy 154 TDIEENPASRYLYMATLYKTYT-----NNVDEFGASFFNKLSFTTGFLFGNGIKRA 202  
Db 76 TRLERFDR----VAQHFRGQYARLINQNNGRLPNSETEISYVKAVIDN---GVSSAA 127

Qy 203 LKQIIRNSLP-----LDIGTE---HSVSRQLHITSSYKDVMDTQTPALPK----- 244  
Db 128 IDLWVINRSLFDMAWDGSYWALGLGIEAERIHNEQANWPNGSERDNRQLISALDKGFDGSF 187

Qy 245 FAKRISLMVWVQRLATVAGYVDPWYK-KW---YMLKFNHVNRPVTFKPFENKE 296  
Db 188 IKGHTTFLQSMMMDVTKLGVEYTIDCWOKIGGGWNGIINDLYKSYKREWVGFELVNN 247

Qy 297 IREPSKALKEVVS-----TDTKDLFENKIGOGTVDFNKEI 332  
Db 248 IKGQNEAKHNEINSLVHDMKAAGKERRGDDLNTOWNLTQAEITVNDIVNTSDCIEKGV 307

Qy 333 RDPSPALKEVSNADKDLFENKIGSGTVDFTINNEIRDPSKALIRKVSTGAEDLFENKIGQ 392  
Db 308 K-AKIELSERKMKNAASLDAGS-----AEKAKQVVEDLAQAAKEAVEN-AK 351

Qy 393 GTVDITFINNEIRDPSKAL-IRKVTEADDLFENKIGQGYTD 431  
Db 352 STAERKAQAAREFFGKLPSRKDLAERFRDIFPNP--EGWID 390

**RESULT 7**

ID Y162\_UREPA STANDARD; PRT; 457 AA.  
AC Q9PQ72;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein\_UU162.  
GN  
OS ureaplasma parvum (ureaplasma urealyticum biotype 1).  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Ureaplasma;  
OX NCBI\_TAXID=134821;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SEROVAR 3;  
RX MEDLINE=2050219; PubMed=11048724;  
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,  
RA Cassell G.H.;  
RT "The complete sequence of the mucosal pathogen Ureaplasma  
urealyticum.";  
RL Nature 407:757-762(2000).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to license@isb-sib.ch).  
CC  
CC EMBL: AE002116; AAC30568.1;  
DR InterPro: IPR00087; Collagen.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 457 AA; 51904 MW; 570B0FCBD04D637 CRC64;

Query Match Score 4.9%; Score 116; DB 1; Length 457;  
Best Local Similarity 19.4%; Pred. No. 1; Mismatches 83; Conservative 61; Mismatches 133; Indels 150; Gaps 19;

Qy 78 LQSKYPLEKIQPTLIPDQVLEAFNLFKES--DANPANSSTEKFRRMFRRGKHSY 134

Db . 25 LFENDQQLERKRKQALT-----TFLRAQSMIDDCVASN-----GAGFIYSN 66  
Oy 135 FHDLYVNLLERKNTRDADA----- 154  
Db 67 FDDKTTDYKTFKKAIAUTNYLWLOGHFRDQYTSAGDGSLSLFKSENADFEKKRL 126

Qy 155 DIENFASRYLYMATLYKTYT-----YTNVDEFGASFFNKLSFTTGFLFGNGIKRA 204  
Db 127 DIN-----KIRVGGLYKKSFRANFNDGIVKVNVDDELKINAIN-----NVHEELL 171

Qy 205 Q-----LIRSNLPLDIGHESVSRQLHITSS-----KRDYM--DTOTP--ALPKFA 246  
Db 172 DHVSKNNPVVNNFEDLN---HHNINNVSVTASGNLATLNGFLNDCTLPNDLTDSN 228

Qy 247 KRFSLMVWVQRLATVAGYVDPWYK-WMFLKFNHVNRPVTFKPFENKEIRPSKAI 304  
Db 229 KVD-----NATANVJUDVKNLAWGQONNNNATR-----QELYEVCKNI 270

Qy 305 KEVKSTDTKLFENKIGQGTVDFFNK-----EIRDPSKALKEVSN- 345  
Db 271 NSKMYNDA---IYDVKAGELVYDDENTNTIKVYKALVNINGVSTHNIVDPSQKIQINNEWL 328

Qy 346 -DAKDLFENKIGQGTVDFTINNEIRDPSKALIRKVSTGAEDLFENKIGQGYTDINNEIRD 404  
Db 329 INIQQDLNLIKKEGVQDAVDNKVNEVKKAIDTKENTFTKML-ENKVNEAKQAVLSSYEQD 387

Qy 405 PSKALIR 411  
Db 388 KASGVFK 394

**RESULT 8**

ID MLP1\_YEAST STANDARD; PRT; 1875 AA.  
AC Q02455;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Myosin-like Protein MLP1.  
GN MLP1 OR YKR0495W OR YKR0415.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycetidae; Saccharomycetes;  
OC Saccharomycetales; Saccharomyctaceae; Saccharomyces.  
OX NCBI\_TAXID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C;  
RX MEDLINE=93247549; PubMed=8483450;  
RA Koellling R., Nairu T., Chen E.Y., Botstein D.;  
RT "A new yeast gene with a myosin-like heptad repeat structure.";  
RL Mol. Gen. Genet. 237:359-369(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94205265; PubMed=8154186;  
RA Bou G., Esteban P.F., Baladrón V., Gonzalez G.A., Cantalejo J.G.;  
RA Remacha M., Jiménez A., del Rey F., Ballesta J.P.G., Revuelta J.L.;  
RT "The complete sequence of a 15,820 bp segment of Saccharomyces  
cerevisiae chromosome XI contains the UBR2 and MLP1 genes and three  
new open reading frames.";  
RT Yeast 9:1349-1354(1993).  
RL Yeast 9:1349-1354(1993).  
CC REPAIR  
CC - FUNCTION: MYOSIN-LIKE PROTEIN THAT IS PROBABLY INVOLVED IN DNA  
CC - SIMILARITY: SOME, TO THE TPR ONCOGENE.  
CC - CAUTION: REF. 2 MISQUOTES THE GENE NAME AS "MLP1".  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to license@isb-sib.ch).

CC  
DR EMBL; L01992; AAA34783.1; -.  
DR EMBL; X73541; CA851948.1; -.  
DR EMBL; 228320; CA82174.1; -.  
DR PIR; S38173; S38173.  
DR SGD; S001803; MLPL.  
KW Coiled coil; DNA repair.  
FT DOMAIN 69 487 COILED COIL (POTENTIAL).  
FT DOMAIN 531 1678 COILED COIL (POTENTIAL).  
FT DOMAIN 1834 1866 COILED COIL (POTENTIAL).  
FT CONFLICT 301 301 R -> A (IN REF. 1).  
SQ SEQUENCE 1875 AA: 218455 MW: 68380D3C9068867 CRC64:  
Query Match 4 9% Score 115; DB 1; Length 1875;  
Best Local Similarity 20.0%; Pred. No. 6-9;  
Matches 109; Conservative 92; Mismatches 211; Indels 134; Gaps 24;  
QY 4 SDPSVGDVTKTLAASEEVDASAAANAYMINSDMSDYLSSVSDNFAERICCSOPKGNSNSAVY 63  
Db 523 SDSQKVYTERLV-----EFKNNIQLOQEKNALLKVKR-NLADKLESKEKKSKOSIQLK 574  
QY 64 SAYMSRCAKQDCLTQSILKYPLEAKVQPLTPDPPYQLEAFALFKESDANPASTERFW 123  
Db 575 ESEVTNEAKEATITLKEKMDLIESVEL--OKEELKLTSVPNDASVNTVKTOL 630  
QY 124 MRFR-----RGKHNSYFHDLVFNILE----- 144  
Db 631 EFKRKLESQVQQLQTRISQITRESTERNSLMLKQIQLDYDSKSDISIKLGKEKSRILAE 690  
QY 145 --KNVTDADATDIEN--FASRYLYMATY---KTYTNVDEFGAFFENLIFT-TGL 194  
Db 691 ERFKLISNTLDLTKAENDOLKRKFDFYLQNTLKDQSKTHETLINEY-VSCKSKSLIVTEL 749  
Qy 195 FGWGIKRAKQIIRSNPLDI-----GTEHVSRLQHITSSYKYM----- 235  
Db 750 LMKKEOKLRVHLJEKNIQELKLSPEKDSLJMLVYQLOIQUKEREDPLEETRKSCQKI 809  
QY 236 DTQQIPALPKFKRFSLM--VVORLLATVAGYVDTPWKKWYMKLKNFMVNRYFTPTKKFF 293  
Db 810 DLELDALSELKETTSQDHHKIQLQLEDDNNSTE-WXQNKIEALKDQYESVITSVDSK-- 865  
QY 294 NREIREPSKALKEVSTIDTKDFENKIGQOGTVDFENKEIROPSS--KAL-KEKV----- 344  
Db 866 ---QTDIEKLQYKVSKLEKEIEDKIRLHTYNMDETFINDLSRKELEKSKNLTDAYS 921  
Qy 345 --NDAKOLF-----NKGQGVDFIN--NEDRPSKALKRVSTGDELHENKI 390  
Db 922 QIKEKYLIYETTSQSLOOTNTIDESKFDTQKNIKNTLEDKISLQREOMEN-- 978  
QY 391 GOCTVDPFINNEIRDPSKALIRKVVTEADDLFENKIGQOGTVDFINKE--IRDPSKALIRK 447  
Db 979 -----LNNELDQKKGMEKE--KAD--FKKRI-SILQNNKEVAEKVEAKSEYESLISK 1024  
QY 448 VSTEAD 453  
Db 1025 IQNDLD 1030

RESULT 9  
MTS2\_DICDI STANDARD; PRT; 2116 AA.

---

CC  
ID MTS2\_DICDI STANDARD; PRT; 2116 AA.  
AC P08799; 01-NOV-1988 (Rel. 09, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Myosin II heavy chain, non muscle.  
GN MICA.  
OS Dictyostelium discoideum (Slime mold).  
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.  
OX NCBI\_TaxID=44689;  
RN [1] SEQUENCE FROM N.A.  
RP MEDLINE-#87092266; Pubmed=3540939;

RX STRAIN=AX2,  
RC "Conserved protein domains in a myosin heavy chain gene from Dictyostelium discoideum.",  
RT Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).  
RN [2] PHOSPHORYLATION SITES, AND MUTAGENESIS.  
RP RC STRAIN=AX2,  
RX MEDLINE=9033583; PubMed=2387408;  
RP RT Lueck-Vielmetter D., Schleicher M., Grabatin B., Wippler J.,  
RA Gerisch G.;  
RT "Replacement of threonine residues by serine and alanine in a phosphorylatable heavy chain fragment of Dictyostelium myosin II.",  
RL FERS Lett. 269:239-243(1990).  
RN [3] PROSPHORYLATION SITES.  
RP RX MEDLINE=8812226; PubMed=2828113;  
RA Waege G., Noegel A., Scheel J., Gerisch G.;  
RA "Phosphorylation of threonine residues on cloned fragments of the Dictyostelium myosin heavy chain.",  
RL FERS Lett. 227:71-75(1988).  
RN [4] X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.  
RP RX MEDLINE=95345066; PubMed=7619795;  
RA Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,  
RA Rayment I.;  
RT "X-ray structures of the myosin motor domain of Dictyostelium discoideum complexed with MgADP Befx and MgADP.AfE4-.",  
RL Biochemistry 34:8960-8972(1995).  
RN [5] X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.  
RP RX MEDLINE=95345067; PubMed=7619796;  
RA Smith C.A., Rayment I.;  
RT "X-ray structure of the magnesium(II)-pyrophosphate complex of the truncated head of Dictyostelium discoideum myosin to 2.7 Å resolution.",  
RT Biochemistry 34:9873-9881(1995).  
RL RN [6] X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.  
RP RX MEDLINE=96206189; PubMed=8611530;  
RA Smith C.A., Rayment I.;  
RT "X-ray structure of the magnesium(II).ADP-vanadate complex of the Dictyostelium discoideum myosin motor domain to 1.9-Å resolution.",  
RL Biochemistry 35:5404-5417(1996).  
RN [7] X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.  
RP RX MEDLINE=97452580; PubMed=9305951;  
RA Guick A.M., Bauer C.B., Thoden J.B., Rayment I.;  
RT "X-ray structures of the Dictyostelium discoideum myosin motor domain.",  
RL Biochemistry 36:11619-11628(1997).  
RN [8] X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.  
RP RX MEDLINE=98070605; PubMed=9405148;  
RA Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;  
RT "X-ray crystal structure and solution fluorescence characterization of Mg<sub>2</sub>(3')-O-(N-methylanthraniloxy) nucleotides bound to the Dictyostelium discoideum myosin motor domain.",  
RL J. Mol. Biol. 274:394-407(1997).  
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE ACTIVITY THAT IS ACTIVATED BY ACTIN.  
CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES INTO FILAMENTS. HEMMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
CC -1- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL CORTEX.  
CC -1- DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMOTIN (LMM) AND 1 HEAVY MEROMOTIN (HMM). IT CAN BE FURTHER SPLIT INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED SUBFRAGMENT (S2).  
CC -1- DOMAIN: THE ROD-LIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES, CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.





FT	CHAIN	346	566	HEMAGGLUTININ HA2 CHAIN.
FT	CARBONYD	27	27	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBONYD	28	28	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBONYD	40	40	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBONYD	104	104	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBONYD	144	144	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBONYD	172	172	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBONYD	177	177	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBONYD	286	286	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBONYD	304	304	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBONYD	498	498	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	SEQUENCE	557	557	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	566 AA:	63514 MW:	AEC60738C765F637 CRC64;
Query Match		4.7%;	Score 112;	DB 1; Length 566;
Best Local Similarity		20.5%;	Pred. No.	2.5;
Matches		93;	Conservative	68;
Mismatches		162;	Indels	130;
Gaps		23;		
Qy	33	DMSDYSIAMS--DNF---AERICCSQVPSKGSNCNSAVASVAMSACRQD-----CLTLOS	80	DR EMBL: Z70650; CA94624.1; -.
Db	117	YFHDLYVNLLEKNVTRADATIDENFAFSRQLYMATLYKTYTNVDEFGASF--FNKLSF	170	KW Hypothetical protein.
Db	81	YVSVVSYNNNRRTPEAERPVRKVSGQRANRFLPEPDITIEANGNLIAPRHAFA	133	SBQUENCE 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;
Db	173	GSYP-----NLSKSYNNKEKEVLVLMGWSVHRSNIEDQTYR---KENA	214	DR EMBL: Z70650; CA94624.1; -.
Qy	134	YFHDLYVNLLEKNVTRADATIDENFAFSRQLYMATLYKTYTNVDEFGASF--FNKLSF	190	DR EMBL: Z70650; CA94624.1; -.
Db	215	YVSVVSYNNNRRTPEAERPVRKVSGQRANRFLPEPDITIEANGNLIAPRHAFA	274	DR EMBL: Z70650; CA94624.1; -.
Qy	191	TTGLFEGWIKRALKQITRSNPLDGTGHESVSRQLHITSS-YKQMDTOIPALPKFAKR	248	DR EMBL: Z70650; CA94624.1; -.
Db	275	NRG-EGSG-----ITTSNAMDECDTKCTOPOGAINSSLPFQNIHPVTIGCPKYRS	326	DR EMBL: Z70650; CA94624.1; -.
Qy	249	FSLMVV-----ORLIATVAGYVDPW--YKKWY-----M	275	DR EMBL: Z70650; CA94624.1; -.
Db	327	TKLRLMTGLRNIPSQTSRGIFGAGTAGFEGGWTGMIDGWGYHVNHOEGSGYADQKSTQ	386	DR EMBL: Z70650; CA94624.1; -.
Qy	276	KLKNFNVNRY-----FIPTKFFENKEIRRSKALKERKVSTDTK-----L	315	DR EMBL: Z70650; CA94624.1; -.
Db	387	NAINGITKVNVSVIEMKNTQFTAVAKKEFD-ELEKRMNENLNKVKVDPGFLDIWTYNAELLV	445	DR EMBL: Z70650; CA94624.1; -.
Qy	316	FENKIGQGTDFPNKEIRDPSKALKEKRSVDAKDFENKIGQGTDF--INNEIRDPSK	372	DR EMBL: Z70650; CA94624.1; -.
Db	446	LEN---ERTLDFHDNSAVKNUYEVKWSQLKNNAE---IGNCCEFHYKCNE-----	491	DR EMBL: Z70650; CA94624.1; -.
Qy	373	ALRKVSTGAEDL---FENKIGQGTDFNNE	401	DR EMBL: Z70650; CA94624.1; -.
Db	492	-CMESVNRGTYDYPKYSSEESKLNRERIDGVLE	523	DR EMBL: Z70650; CA94624.1; -.
RESULT	12			
YD86_SCPO				
ID	YD86_SCPO			
AC	010411;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DE	Hypothetical 222.8 kDa protein C1F3_06C in chromosome I.			
GN	SPAC1F3_06C.			
OS	Schizosaccharomyces pombe (Fission Yeast).			
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;			
OC	Schizosaccharomycotales; Schizosaccharomycetaceae;			
OC	Schizosaccharomyces.			
OX	NCBI_TAXID=4896;			
RN	[1]			
RP	SEQUENCE FROM N. A.			
RC	STRAIN=972;			
RA	Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;			
RL	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
CC	This SWISS-PROT entry is copyrighted. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> )			
CC	or send an email to license@isb-sib.ch).			
CC	.....			
CC	DR EMBL: Z70650; CA94624.1; -.			
CC	KW Hypothetical protein.			
CC	SBQUENCE 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;			
Qy	9	DVTKLIAASISVSDSANAY-MINSDMSDYL-SAVSDNFA-----ERICCSQVPKSN	58	DR EMBL: Z70650; CA94624.1; -.
Db	205	DLSRQILTVTEKLDKKEKDVKEDVSSIKSLAEEQANSKSLURGEQESEL-EKLLVSS	263	DR EMBL: Z70650; CA94624.1; -.
Qy	59	CSAVSAYMSRCAKODCLQ-----SKYPLAKYQOLTLPPQYQLE	101	DR EMBL: Z70650; CA94624.1; -.
Db	264	KTVSTLROTEENSLRAECKTQDKLEKCAINEBEDSKLLEELKHNV-ANYSAIVHKDKLIE	322	DR EMBL: Z70650; CA94624.1; -.
Qy	102	AAFLIKFESDANPANSTEKKWRMRFRGK-----NHSYFHDLVNLIE	144	DR EMBL: Z70650; CA94624.1; -.
Db	323	DLISTRISEFD---NLKSEROTLSTNEKEKLRLRTGSLKDSRTSNSOLEEENVELKE	378	DR EMBL: Z70650; CA94624.1; -.
Qy	145	KNVTRDADATIDENFAFSRQLYMATLYKTYTNVDEFGASF--FNKLSF	190	DR EMBL: Z70650; CA94624.1; -.
Db	379	SNRTHSQLDAESKSF--EQENKSLKSIDEXQ---NNLSSKDKM---vQVSS	427	DR EMBL: Z70650; CA94624.1; -.
Qy	205	QI---INSNLPLDGTGHESVSRQLHITSS-YKQMDTOIPALPKFAK--REFLMLVYQRLAT	260	DR EMBL: Z70650; CA94624.1; -.
Db	428	QLEARSSL---AHATGKLAETNSE-RDFQNKKIDFPEKIEODLRACQINSSNELKE	480	DR EMBL: Z70650; CA94624.1; -.
Qy	261	VAGYVTPWPKWMLKFNVNRYFIPFKENETREPSKALKEKVSTDTK--DLEEN	318	DR EMBL: Z70650; CA94624.1; -.
Db	481	KSALID---KKDQELNL--REQIKEKVSESTOSSQSLORDILNEKKHEVES	532	DR EMBL: Z70650; CA94624.1; -.
Qy	319	KI---GOGTVDFENKE-----RIGQTVDFINN-----ERDPSKALKEKVSN	345	DR EMBL: Z70650; CA94624.1; -.
Db	533	QNLNLKGELQETISNESEHLSQLSLOTLAAKEEAATNNELSESKEQNLSQTLCNAFOEKLAK	592	DR EMBL: Z70650; CA94624.1; -.
Qy	346	DAKDIFEN-----RIGQTVDFINN-----ERDPSKALKEKVSTGADEFLE	387	DR EMBL: Z70650; CA94624.1; -.
Db	593	SVMQLKENEQNQFSSLTSFKKLNEQSHOLENNHQITQKDTSKLQQLQERLAN-FE	650	DR EMBL: Z70650; CA94624.1; -.
Qy	388	NKIGQTVDFINNEIR-----DPSKALKEVYFADDIFENKIGQGTDFINKEIRD	440	DR EMBL: Z70650; CA94624.1; -.
Db	651	OK---ESTLSDENNDLRLKLEESNLSLRK-QEDVDSELKN-----TQTLKEDLRK	701	DR EMBL: Z70650; CA94624.1; -.
Qy	441	SKALIRKVSTAEIDLLE	457	DR EMBL: Z70650; CA94624.1; -.
Db	702	EEAL-RFSKLEAKNRE	717	DR EMBL: Z70650; CA94624.1; -.
RESULT	13			
ALP6_SCPO				
ID	ALP6_SCPO			
AC	010412;			
DT	01-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DE	Spindle pole body component Alp6.			
GN	ALP6 OR SPBC42.01C OR SPBC42.20C.			
OS	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;			
OC	Schizosaccharomyces pombe (Fission Yeast).			
OC	Schizosaccharomycotales; Schizosaccharomycetaceae;			
OC	Schizosaccharomyces.			
OX	NCBI_TAXID=4895;			
RA	[1]			
RL	SEQUENCE FROM N. A., AND CHARACTERIZATION.			
CC	MEDLINE=2053253; Pubmed=11080156;			
CC	Vardy L., Toda T.;			

RT \*The fission yeast gamma-tubulin complex is required in G(1) phase and  
 is a component of the spindle assembly checkpoint.;  
 RL P54788;  
 RN EMBJ J. 19:6098-6111(2000).  
 RP [2]  
 SEQUENCE OF 1-566 FROM N.A.  
 RC STRAINE-972;  
 RA Saunders D., Harris D., McDougall R.C., Rajandream M.A., Barrell B.G.;  
 RL Submitted (Nov-1999) to the EMBL/GenBank/DDBJ databases.  
 CC -1 FUNCTION: COMPONENT OF THE GAMMA TUBULE COMPLEX THAT IS REQUIRED  
 FOR THE REGULATION OF BOTH INTERPHASE MICROTUBULES AND MITOTIC  
 BIPOLEAR SPINDLES.  
 CC -1 SUBCELLULAR LOCATION: SPINDLE POLE BODY AND THE MICROTUBULE  
 ORGANIZING CENTER (MTOC).  
 CC -1 SIMILARITY: BELONGS TO THE SPB ALP6/SPC98 FAMILY.  
 CC -1 CAUTION: REF 1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A  
 CC FRAMESHIFT IN POSITION 1.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 or send an email to license@isb-sib.ch).  
 CC -1 SIMILARITY: BELONGS TO THE SPB ALP6/SPC98 FAMILY.  
 DR EMBL; AB040011; BAA94097.1; ALT\_FRAME.  
 DR EMBL; AL133306; CAB62095.1; .  
 KW Microtubules; Mitosis;  
 SQ SEQUENCE 832 AA; 95996 MW; C92271C3DBF5C01A CRC64;

Query Match 4.7%; Score 111.5; DB 1; Length 832;  
 Best Local Similarity 20.0%; Pred. No. 4.3; Mismatches 194; Indels 129; Gaps 23;  
 Matches 101; Conservative 80;

Qy 4 SDSVGDYTKTLAASESDSAANAYMINSDMSDYLSDAVSDNPAERICSQVKG----- 56  
 Db 107 SOSSRDSVSSHLL--DESISNPINPSTEVE-----S-SNGQTRYDQVPEMPQTDDW 156  
 Qy 57 ---SNCASASVASYMSRCAQ-----DCLTLQSLSKYPLEAKYQPLTPDPYQL 100  
 Db 157 EGLENESIISTANDSSLRNLRSTETSSVQHTLITEADALISIISVLOE-----I 204  
 Qy 101 EAFILEKESDANPANSTEKRWMRFRP--GKHNISFYHDL-VNLLEKNVTROADATDIE- 157  
 Db 205 STEYVQSKNELLASKRIPVQQLQMRSALSERGFLYQELKVFSNYDVSQSIDGVNSK 264  
 Qy 158 --NFASRYLV-MATLYKYTYNDEFGASFFNKLSTFTGL-----PGWGIKRAL 203  
 Db 265 AFINDOSLALQSLSIKSVISKELTNFLAIIASLDQSQRADASLSERKPMVITRCA-TAWTQVAKL 324  
 Qy 204 KOIRSNLPLDGTGERSVSRQHITSSYKDYMDTQIPALPKFAKRFSLMMVORLATVG 263  
 Db 325 KLRILSSVNDNNQENKKRLIQVSKVNVHGDPLQELS-----DKILTEITG 373  
 Qy 264 YVDTPWK--KIVWKMUKKFNWNRVFLPTKKFENKE TREPSFALKEVSTTDKLFENKK 320  
 Db 374 --PLXEMIENWIYK----GELVQPYQERF-----VKEKNGSESHD---HQ 409  
 Qy 321 GOGTV----DFENKEITRD--SKALLEKVKSDAKDLFENKIGOTVDFINNEIRDPSKA 373  
 Db 410 GOGDVWVKKGKFIDKELLIPLPSFLSEELVVDKIFLIGSKINFARYGCGDFTW---AQENYOK 465  
 Qy 374 LIRKVSTGADEFENK1GQTVDFINNEIRDPSKALTRKVYEAADDIENK---IGOT 429  
 Db 466 IIVKLSYFDPHSELETWVTDKAYEESINLV-----VLMEEFHLDHKIAIKYLLLGQG- 519  
 Qy 430 VDFIN----REIOPSKALTR 446  
 Db 520 -DFVDLIMESLGNSDLQDPAINTLFR 542

CC SEQUENCE FROM N.A.  
 DR RX MEDLINE-96099401; PubMed=7502077;  
 RA Gavin K.A., Hidaka M., Stillman B.;  
 RL "Conserved initiator proteins in eukaryotes.";  
 CC Science 270:1667-1671(1995).  
 CC -1 FUNCTION: COMPONENT OF THE ORIGIN RECOGNITION COMPLEX (ORC) THAT  
 BINDS ORIGINS OF REPLICATION. IT HAS A ROLE IN BOTH CHROMOSOMAL  
 REPLICATION AND MATING TYPE TRANSCRIPTIONAL SILENCING. BINDS TO  
 THE ARS CONSENSUS SEQUENCE (ACS) OF ORIGINS OF REPLICATION IN AN  
 ATP-DEPENDENT MANNER.  
 CC -1 SUBUNIT: ORC IS COMPOSED OF SIX SUBUNITS (BY SIMILARITY).  
 CC -1 SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1 SIMILARITY: BELONGS TO THE ORC1/CDC6/CDC18 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 or send an email to license@isb-sib.ch).  
 DR U40151; AAC9130.1; .  
 DR InterPro; IPR003950; AAA\_subfam.  
 DR IPRO01025; BAH.  
 DR PF00004; AAA; 1;  
 DR Pfam; PF01426; BAH; 1.  
 DR SMART; SM00439; BAH; 1.  
 KW DNA replication; Nuclear protein; DNA-binding; ATP-binding.  
 FT NP\_BIND 471 478 ATP (POTENTIAL).  
 SQ SEQUENCE 886 AA; 101393 MW; 2DA665BC291F182F CRC64;

Query Match 4.7%; Score 110.5; DB 1; Length 886;  
 Best Local Similarity 19.5%; Pred. No. 5.4; Mismatches 198; Indels 129; Gaps 23;  
 Matches 99; Conservative 82; Mismatches 198; Indels 129; Gaps 23;

Qy 5 DSYGDYTKTLAASESDSAANAYMINSDMSDYLSDAVSDNPAERICSQV----- 53  
 Db 274 EAISNEDSLSDEYHESEKKEFANASSSDSDEFEDYOSAELATIVEPAKKVVRKSPKDPI 333  
 Qy 54 --PKGSNCASASVASYMS--RCAODCLTLQSLSKYPLEAKYQ-PLTPDPYQLEAFILEK 108  
 Db 334 SPVKSTQTPLOPSAVHSPPRFKKNNIVRAKAVYTFPSKRYKNPKDIPDLD----IFQR 368  
 Qy 109 ESDANPANSTEKRWMRFRGKHNISFYHDLVLNILEKVNTRD ADDATIENF ASRYLYM 166  
 Db 389 HNNIDLIAALEERFRTVSAKGKMETIFSKVKQKLNRSKEEVKAADFNYLPAREN 448  
 Qy 167 ATLYKTYVNDDEGASFFNKLSTFTGLFGWGKIRKALKOITRSNPLDGT---EHSYFSR 223  
 Db 449 ASTYLSLSAI-EAGTSTSIYIAGTPGV--GKTFITVREVK---DLMTSADQKELPR 499  
 Qy 224 LOHT-----TSYK-----DMDTQIPALPKFKRSLMVOR 256  
 Db 500 FQYIEINGJKIVKASDSTEVFWQKISGEKLTSGAMESLEFYFVKVPAKTKRPIVVLDE 559  
 Qy 257 LLATVAGVYDTPW-YKKW-YMKLKNFMY--NRFVFPKKEPNK----- 295  
 Db 560 LDALVSKSDQVMNFNFNWATYSNAKLVIVAVAVANTLDLPERHLGNKISSRGTRIMFTGY 619

RESULT 14  
 ORCL\_KLUIA



RESULT 16  
**HEMA\_JAPUE** STANDARD; PRT; 566 AA.  
ID HEMA\_JAPUE STANDARD; PRT; 566 AA.  
AC P03452; 083964;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DE Hemagglutinin precursor [Contains: Hemagglutinin HA chain; DE Hemagglutinin HA2 chain].  
GN HA  
OS Influenza A virus (strain A/Puerto Rico/8/34).  
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;  
OC Influenza virus A and B group; Influenza A viruses;  
OC NCBITaxonID=11455;  
RN [1]  
RP SEQUENCE FROM N.A.; MEDLINE=8213600; PubMed=7278958;  
RX RA Winter G., Fields S., Brownlee G.G.;  
RT "Nucleotide sequence of the haemagglutinin gene of a human influenza virus H1 subtype.";  
RT Nature 292:72-75(1981).  
RN [2]  
RP SEQUENCE OF 18-243 FROM N.A.  
RT STRAIN="SUBTYPE H1"; MEDLINE=83129355; PubMed=6186384;  
RA Cation A.J., Brownlee G.G., Yewell J.W., Gerhard W.;  
RT "The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype)."; Cell 31:417-427(1982).  
RT -1. FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO CELL RECEPTORS AND FOR INITIATING INFECTION.  
-1. SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS (-1 HAL AND HA2) LINKED BY A DISULFIDE BOND.  
-1. SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: V01088; CAA24272.1; -.  
DR EMBL; J02144; AAA43194.1; -.  
DR PIR; A04063; HMV.  
DR HSSP; P03437; 5HMG.  
DR InterPro; IPR001364; Hemagglnin.  
DR Pfam; PF00509; Hemagglnin\_1.  
DR PRINTS; PRO00329; HEMAGGLUTIN12.  
DR PRODOM; PD000225; Hemagglnin\_1.  
DR KWRP; A04063; HMV.  
DR SIGNAL; 17  
FT CHAIN 18 343 HEMAGGLUTININ HA1 CHAIN.  
FT CHAIN 345 565 HEMAGGLUTININ HA2 CHAIN.  
FT CARBOHYD 27 27 N-LINKED (GlcNAc. . .) (POTENTIAL).  
FT CARBOHYD 28 28 N-LINKED (GlcNAc. . .) (POTENTIAL).  
FT CARBOHYD 40 40 N-LINKED (GlcNAc. . .) (POTENTIAL).

RESULT 17  
**TIRL\_SCCHO** STANDARD; PRT; 577 AA.  
ID TIRL\_SCCHO  
AC Q09684;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Nuclear fusion protein tth1.  
GN TIRL OR SPAC13C5.03  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomyces.  
OX NCBITaxonID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.; MEDLINE=98106170; PubMed=9442101;  
RX Tange Y., Horio T., Shimamaki M., Ding D.-Q., Hirakawa Y., Niwa O.;  
RT "A novel fission yeast gene, tth1+, is required for the fusion of nuclear envelopes during karyogamy.";  
RL J. Cell Biol. 110:247-258(1998).  
RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA Devin K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;  
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- PTM: N-GLYCOSYLATED.

CC -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----  
 EMBL: D07337; BAA1334 1; -  
 DR EMBL; Z50112; CAA90454 1; ALT\_INIT.  
 DR Transmembrane; Glycoprotein.  
 KW TRANSMEM 405 425  
 FT TRANSMEM 434 454  
 FT CARBOHYD 163 163  
 FT CARBOHYD 372 372  
 SEQUENCE 577 AA; 66572 MW; 717DDC99AE45EF CRC64;

Query Match 4.7%; Score 110; DB 1; Length 577;  
 Best Local Similarity 19.2%; Pred. No. 3.4; Gaps 24;  
 Matches 94; Conservative 76; Mismatches 139; Indels 180; Gaps 24;

OY 4 SDSYGDVYKTL---LAASESV---DSANAYMINNSDMSDYLSAVSDNFAERICSQVPKG 56  
 Db 25 SESTGVDVSEPMKVVNVAFSEGLAGFDLSAHVY-----QALLR 62  
 Qy 57 SNCSSASVSA-YMSCAKQDCLTQSLSKPLEAKYQPLTLDPYOLEAFLF-----K 108  
 Db 63 STCYOEVAAATLISKCSLN---TELTIDNRIRHSQ-MTLCDF---FERSQILAPSECVRGS 116

Oy 109 ESDKANPANSTEKKRWRERRGKHNISFFHDL---VFNL-LEAKNTRADATIENFRSR 162  
 Db 117 QSECVSKLESTSWLSTP---SHFDVNHLCLRLANEMQRELSTEVN-MWTVLQKQ 170

Oy 163 YLYMATLVKYTYTNV-----DEFGASFFPNKL-SFTGFLFGWNCIKRALQQTIRSNPLD 214  
 Db 171 FLERIVILHLRNPFESTDKNQRIDKDFGKF-----200

Oy 215 IGPPIHVSLSLOHTTSYKVMDQIPALPKFAKRFSLMVQRLATVAGVWDTPPWKKWY 274  
 Db 201 ---NSV-----TENSFKD-----211

Oy 275 MKLKNFMVNRVETPKKFN-----KEIRPSKALEKYSITDKLFENWKIGOTV 325  
 Db 212 ---NFRVNOEIQMGLVELONHQBGWVOKEILSTIKOLKSEI-FDINSFFAN-----259

Oy 326 DFFNKKEIRDPSKALKEKYSINDAKDLFENWKIGOTVDFINNEIRDPSKALIRKVSTGAEDL 385  
 Db 260 --FTEESACGYSNSLJEKINE--EIRDPSKALIRKVTEADDLDFENWKIGOTVDFEINK 315  
 Oy 386 FENKIGKDFEINN-----EIRDPSKALIRKVTEADDLDFENWKIGOTVDFEINK 435  
 Db 316 LQGSV-OSNIDFWNSGQFPLYDLTIOLKBELOS-KRUSSEOOQLQHEQIOLQWSDFLV 373

Oy 436 EIRPSKAL 444  
 Db 374 S-KDHLKVL 381

RESULT 18  
 ID Y338\_MYCGE STANDARD: PRT: 1271 AA.  
 AC P47580;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical lipoprotein MG338 precursor.

GN MG338.  
 OS Mycoplasma genitalium.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 OC Mycoplasmataceae; Mycoplasma.  
 OC NCBI\_TaxID=2097;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33330 / G-37;  
 RX MEDLINE=96020346; PUBMED=7569993;  
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A., Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M., Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S., Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;

RA "The minimal gene complement of Mycoplasma genitalium.";  
 RT Science 270:397-403(1995).  
 RN [2]

RP SEQUENCE OF 1123-1114 FROM N.A.  
 RC STRAIN=ATCC 33530 / G-37;  
 RX MEDLINE=9407230; PUBMED=8253680;  
 RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;  
 RT "A survey of the Mycoplasma genitalium genome by using random sequencing.";  
 J. Bacteriol. 175:7918-7930(1993).  
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (POTENTIAL).  
 CC -----  
 CC -1- SIMILARITY: BELONGS TO THE MG307 / MG338 FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----  
 DR U30715; AAC1563 1; -  
 DR EMBL; U01809; AAD12341 1; -  
 DR TIGR; MG338; -  
 DR PROTE; PS00013; PROKAR\_LIPOPROTEIN; 1.  
 KW Hypothetical protein; Lipoprotein; Membrane; Signal; Complete proteome.  
 FT SIGNAL 1 26  
 FT CHAIN 27 1271  
 FT LIPID 27 N-ACYL DIGLICERIDE (POTENTIAL).  
 SQ SEQUENCE 1271 AA; 142492 MW; FCE6042067310A70 CRC64;

Query Match 4.7%; Score 110; DB 1; Length 1271;  
 Best Local Similarity 17.7%; Pred. No. 9.1;  
 Matches 114; Conservative 77; Mismatches 220; Indels 232; Gaps 26;

Oy 12 KTLLAASESVDSAANAYMINNSDM---SDYLSAVSDNFA---ERICSQVPKGNSCASYSA 65  
 Db 389 KSAMKSKQSVTSDDNGFNVKSEELKINPSLSSGSDNNTSQFWKQQLANNSQPTI 448

Oy 66 Y---MSRCQDCLTQSLSKPLEAKYQPLTLDPD-----YOLEAFLFRES 110  
 Db 449 FDAVRMESNSQSAQVTSNLVSLSSKTTQQQPKVVRGDDAIYAFHIDGNYFLS 508

Oy 111 DANPANSTKR - -WMRERRGKHSYEDLV-----139

Db 509 SPNKRKRFEKAQEVILMFLQGQTNNFSKDNVFSVDFGSNSFRSWANRNTLKVIAL 568

Oy 140 FNLEEKVTRRADATDIEFASRYLIMATLYKTYTINDEFGASF-----FN 186  
 Db 569 TTMCENGTSNNNGKDVCLAKKJNNTNLSETIKQODDENLSQIKSSYESYIKAN 628

Oy 187 KL-SFTTGL-----PGWG-----198  
 Db 629 KLNNTFIDLANIEKLEQATVDRANNYIKLQEAKESSIGGWGQPLPYKRANDGSYPSLAKF 688

Oy 199 -----IKRAKQIIRSNI-----LDIGTEHSVSL-----QH 226



RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 RT Jannaschii."  
 RL Science 273:; 1058-1073(1996).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation.  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).  
 CC  
 EMBL: U67557; AAB9156.1; -. DR HSSP\_P03437; 2HMG;  
 DR TIGR; MJ147; -. PIR; IPR01364; Hemagglutin.  
 KW HYPOTHETICAL PROTEIN; Transmembrane; Complete proteome.  
 FT TRANSMEM 381 401 PIR; A29971; HMIVS.  
 FT SEQUENCE 462 AA; 54433 MW: 91FF068862971776 CRC64;  
 SQ Hemagglutinin; Envelope protein; Glycoprotein.  
 FT NON\_TER 1 1  
 CHAIN 1 328 HEMAGGLUTININ HA1 CHAIN.  
 FT CARBOHYD 330 550 HEMAGGLUTININ HA2 CHAIN.  
 FT CARBOHYD 8 8 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 550 AA; 61580 MW: 991F6DBBC02F24F2 CRC64;  
 Query Match Similarity 4.6%; Score 107.5; DB 1; Length 462;  
 Best Local Similarity 22.7%; Pred. No. 3.8; Mismatches 75; Indels 45; Gaps 7;  
 Matches 44; Conservative 30; Mismatches 75; Indels 45; Gaps 7;  
 Qy 130 KNHSYFHQDVLVNLLEKVNTRDAD---AVDIEFASRYLYMATLYKKYTVDNEFGASF 184  
 Db 195 KSHYFEDTSFLNSLFTDFRDRVSFKIDKENFKGLAIFLKIQFLNKHTTENFQLS- 253  
 Qy 185 FNKLSPTTGLFGWQIKRALKOITISNLPLDGTG---HSVSRQHHTSYKDMDTOIP 240  
 Db 254 --KSP-----LKTDVKEVLCDVKKEBILSKHLRKEDSNKDKDENLK 299  
 Qy 241 ALPKFAKRFKSLMVQRLATVAGVDTPTYKWWKMLKLFMVF---NRVIFPKFFNKE 266  
 Db 300 ELLGSVY-----1IDQFNYSKSVLAEVRNLRISENNTYEANSLYME 345  
 Qy 297 --TREPYSKALEK 307  
 Db 345 MELTIKFNSNGFEK 359  
 RESULT 2.1  
 HEMALIAZH3 ID HEMA\_IAZH3 STANDARD; PRT; 550 AA.  
 AC P11134; 0844025; Q8B1P6;  
 DT 01-JUL-1989 (Rel. 1.1, Created)  
 DT 01-JUL-1989 (Rel. 1.1, Last sequence update)  
 DT 16-OCT-2001 (Rel. 4.0, Last annotation update)  
 DE Hemagglutinin [Contains: Hemagglutinin\_HA1 chain; Hemagglutinin\_HA2  
 chain] (Fragment).  
 HA Influenza A virus (strain A/swine/Hong Kong/126/82);  
 OS Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;  
 OC Influenza virus A and B group; influenza A viruses;  
 OC Influenza A virus;  
 OX NCBI\_TAXID:11498;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=&8101364; PUBMED=3336940;  
 RT "Origin of the hemagglutinin gene of H3N2 influenza viruses from pigs  
 in China,"  
 RL Virology 162:160-166(1988).  
 CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO  
 CC CELL RECEPTORS AND FOR INITIATING INFECTION.  
 CC -!- SUBUNIT: HOMOTIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS  
 CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND.  
 CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.  
 CC  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).  
 CC  
 DR IntzPO; IPR01364; Hemagglutin.  
 DR Pfam; PF00509; Hemagglutinin; 1.  
 DR Probom; PD000225; Hemagglutin; 1.  
 KW Hemagglutinin; Envelope protein; Glycoprotein.  
 FT NON\_TER 1 1  
 CHAIN 1 328 HEMAGGLUTININ HA1 CHAIN.  
 FT CARBOHYD 330 550 HEMAGGLUTININ HA2 CHAIN.  
 FT CARBOHYD 8 8 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 550 AA; 61580 MW: 991F6DBBC02F24F2 CRC64;  
 Query Match Similarity 4.5%; Score 107; DB 1; Length 550;  
 Best Local Similarity 18.4%; Pred. No. 5.1; Mismatches 164; Indels 194; Gaps 23;  
 Matches 95; Conservative 67; Mismatches 164; Indels 194; Gaps 23;  
 Qy 5 DSGDVTKULLAKESVDSVANAVMINSDMSYLS-----AVSD 43  
 Db 77 DVQEQDETWDLFLVERSNAFSSCSYV---DVPDYASLRSLVASSGTLFITEGFWWTGVIO 132  
 Qy 44 NFAERICSQPKGSNCASASVAYSRSRCAKQCLTQSUKPLEAKYQVQTLPPYQLEA 103  
 Db 133 NGGSACKRG-----ASGFPSRL--NWLUHKSGSPV---LNVMPRNNDNEKL 177  
 Qy 104 FI-----LKFESDANPANSTEK-----RWWMRFRRGKHNHYFHL 138  
 Db 178 YIWGVHHPSTNQEQTNLVQASGRIVTSTRSQTLIPIGSRWVWRGOSGRSIW-- 234  
 Qy 139 VFNLLKEVNTRDADTIEFASRYLYMATLYKKYTVDNEFGASFNLKSFPTGLCGWG 198  
 Db 235 -----TIVKGPDVLVINSGN--LTAAPRGYFKMRT----GRS----- 265  
 Qy 199 IKRAKQIQRISNLPLDGTESVSRVQHIT-----YKDMTDQTPALPKERKSLM----- 253  
 Db 266 -----SINHSDAPIDTCVSECITPNGSFTPNDKPFQNYNKITVGACPKVVKONSLKATG 319  
 Qy 254 -----VORLLATVAGVYDTPW---YKKWY-----MKLK-----NF 280  
 Db 320 MRNNUPEKQKRGFLFGAFLAGIENGKENGMSMIDQWYGRHONSEGTQOAKDKSSTOAIDQING 379  
 Qy 281 MNRYFIPKFFN---KIRREPSKALK--EKVSTDTKOLFENKIGOGIVDFENKEIRD 335  
 Db 380 KLNRVIEKTFNKHQIEKESEVERGIGDQLEKYVEDTK---IDLWS----- 422  
 Qy 336 SKALKERVKPSDAKOLEENIGQGVDFINNEIRPSKALLKIRKSTGADEFLENKIGQTV 395  
 Db 423 -----VNDVFLVALEN---QHRTDLDSEMNKLFEKRROLREADM---GNCF 467  
 Qy 396 DFINNEIRDPDKSKALLRKVYTEDDLFENKIGQGVDF 431  
 Db 468 -----KIKHCDNACIESIRNGCYD 487  
 RESULT 2.2  
 RRPL\_DUGBV ID RRPL\_DUGBV STANDARD; PRT; 4036 AA.  
 AC Q66631; DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 4.0, Last annotation update)  
 DE RNA-directed RNA polymerase (EC 2.7.7.48) (L protein).  
 L. GN  
 OS Dugbe virus.





Query Match 4.5%; Score 106; DB 1; Length 624;  
Best Local Similarity 25.6%; Pred. No. 6.9;  
Matches 68; Conservative 43; Mismatches 87; Indels 68; Gaps 17;

Qy 171 YKTYTNTVDEFGA--S-PFENKLISFTGFLFGWGIKRALKOIIRNSNLPLDIGHESVS-RLOH 226  
Db 321 VINYTGEEEDFSAMATKEPALLCAGSGTELMEKGRREPPAI--NPDVNNEKFISGYLINE 377

Qy 227 ITSSYKDYMTQIPALPK---FAKRFSLMVQRLATVAGYVDPWKKWYMKLNFM 281  
Db 378 LSKPYKE--NTSI-SFPKKNSILKQRFMLMPKSII-----WIMKK---L 417

Qy 282 VNRVFPTKFFENKEI--REPASKALEKVKSTDKDIFENKGQ-----GTVIDFKN 330  
Db 418 VQGVFGFRDKNDEDIPNDPSKMLK--ITKTNSL-NNSAGHKDEIELELLNTSDEYS 473

Qy 331 ETRPSKALKKEVNSDAKDLFENKIGOGTVDFINNEIROPSKALIRKVSTGAEDLENKI 390  
Db 474 EDYEPSE-VEGLGDSDEBNLERDSL-----IFNEFRD--ALL-----DISSSED 514

Qy 391 GOVTVDFINNEIROPSKALIRKVTE 416  
Db 515 NEVHTDY--NNWNSTSRILQQLLSD 538

RESULT 26

ID	HEMA_TADHK	STANDARD:	PRT:	550 AA.
AC	P43257;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Hemagglutinin precursor [Contains: Hemagglutinin HA2 chain] (Fragment).			
GN	HA.			
OS	Influenza A virus (strain A/Duck/Hong Kong/7/75).			
OC	Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;			
OC	Influenza virus A and B group; Influenza A viruses;			
OC	NCBI_TaxID=11364;			
RN	[1]			
RP	SEQUENCE FROM N.A. MEDLINE=91341491; PubMed=1875195;			
RX	Yasuda J., Shortridge K.F., Shimizu Y., Kida H.;			
RT	*Molecular evidence for a role of domestic ducks in the introduction of avian H3 influenza viruses to pigs in southern China, where the RT A/Hong Kong/68 (H3N2) strain emerged.;			
RL	J. Gen. Virol. 72: 2007-2010(1991).			
CC	-1- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO CELL RECEPTORS AND FOR INITIATING INFECTION.			
CC	-1- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS			
CC	(HA1 AND HA2) LINKED BY A DISULFIDE BOND.			
CC	-1- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.			
CC	This SWISS PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).			
CC	EMBL: D00929; BA00769.1; -.			
DR	HSPP; P03437; 5HNG.			
DR	InterPro: IPR01364; Hemagglutin.			
PFam: PF00109; Hemagglutin; 1.				
PRODom: PD000225; Hemagglutin; 1.				
DR				

Query Match 4.5%; Score 105.5; DB 1; Length 550;  
Best Local Similarity 18.0%; Pred. No. 6.4;  
Matches 88; Conservative 69; Mismatches 154; Indels 177; Gaps 22;

Qy 12 KTLLASBESVDSAANAVMINSDMSDYLSAVSINFAERICSCQPKGNCASAVSYASRSRA 71  
Db 109 RLVASSGTLETFTEGIFT-----WIVGTTGGSACKRGP-----ASGFRSL- 151

Qy 72 KDCDLTQSLSKPLEAKYQPLTLDPQLEAFI-----LFKESDANPANS 117  
Db 152 - NWLTLKSGSTPV----LNVTMPNNDNFKDLYIWGVHPRSTNQEQTLYVQASGRVTS 205

Qy 118 TSK-----REWMRFRGKHNHSYFHDLVNLLEKVNTRDADATDIEFASRYLM 166  
Db 206 TRRSQQTIIPNIGSPRWGQSGRSRISW-----TIVKPGDVLVINSGN--LIA 253

Qy 167 ATLYYYKTYTNTVDEFGASPFENKLISFTGFLFGWGIKRALKOIIRNSNLPLDIGHESVS-RLOH 226  
Db 254 PROVKART-----GK-----SIMKSDAPIDTCVSECIPNGS 287

Qy 227 ITSS-YKDYMTQIPALPKFAKRFSLMV-----VORLATVAGYVDPN--Y 270  
Db 288 IPNDKPFQONVNRYTGYACPKVYKQNSLKLATGMRNPEKOTRGLEFAITENGWEGMI 347

Qy 271 KWKW-----MLKL-----NEWNRVFIPTKFFFN--KEIREPSKALK 305  
Db 348 DSWYGFRRHQNSEGTGQAIDLSTQAIQDQINGKLNRVKIKKNEKPHOIKEKFSEVGRIQ 407

Qy 306 - EKVSTDTKDLFENKIGOGTVDFENNEIROPSKALRKVSTGAEDLENKI 363  
Db 408 DLKRYVEDTK-----IDLWS-----YNAVLVALEN--QHTIDL 440

Qy 364 NNEIRDPSKALIRKVSTGAEDLENKI 407  
Db 441 DSEMNLKFEKTRRQLRENAADM----GNGCF-----KIVHKCDNACIE 479

Qy 424 KIGQGTVD 431  
Db 480 SIRNTYD 487

RESULT 27

ID	HEMA_TAE7	STANDARD:	PRT:	566 AA.
AC	P03440;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Hemagglutinin precursor [Contains: Hemagglutinin HA1 chain; DE Hemagglutinin HA2 chain].			
GN	HA.			
OS	Influenza A virus (strain A/England/321/77).			
OC	Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;			
OC	Influenza virus A and B group; Influenza A viruses;			
OC	Influenza A virus.			
OX	NCBI_TaxID=11378;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=83110955; PubMed=6822816;			
RA	Hauptmann R., Clarke L.D., Mountford R.C., Bachmayer H., Almond J.W.;			
RT	"Nucleotide sequence of the haemagglutinin gene of influenza virus			



Db	424	LIEBLHQKDTYVFDLQK	442
RESULT	29		
QAQS_SCHPO	STANDARD;	PRT;	2670 AA.
QAQS_SCHPO	010105;		
DT	01-FEB-1996 (Rel. 33, Created)		
DT	01-FEB-1996 (Rel. 33, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Potentiative translational activator C18G6.05C (GCN1 homolog).		
GN	SPAC18G6.05C.		
OS	Schizosaccharomyces pombe (Fission yeast).		
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;		
OC	Schizosaccharomyctales; Schizosaccharomyctaceae;		
OX	Schizosaccharomyces.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=972;		
RA	Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.:		
RL	Submitted (DCC-1995) to the EMBL/Genbank/DBJ databases.		
CC	-1-SIMILARITY: CONTAINS 19 HEAT REPEATS.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> ) or send an email to license@isb-sib.ch).		
CC	EMBL: 268198; CAA92385.1; -		
DR	InterPro: IPR000357; HEAT_repeat.		
DR	PIFM; PPF02985; HEAT; 4.		
DR	PROSITE: PS00077; HEAT_REPEAT; 4.		
KW	Hypothetical protein; Translation regulation; Activator; Repeat..		
FT	REPEAT 315 352 HEAT 1.		
FT	REPEAT 1062 1099 HEAT 2.		
FT	REPEAT 1319 1356 HEAT 3.		
FT	REPEAT 1439 1476 HEAT 4.		
FT	REPEAT 1478 1514 HEAT 5.		
FT	REPEAT 1518 1555 HEAT 6.		
FT	REPEAT 1557 1593 HEAT 7.		
FT	REPEAT 1637 1674 HEAT 8.		
FT	REPEAT 1676 1713 HEAT 9.		
FT	REPEAT 1714 1751 HEAT 10.		
FT	REPEAT 1755 1792 HEAT 11.		
FT	REPEAT 1793 1830 HEAT 12.		
FT	REPEAT 1898 1939 HEAT 13.		
FT	REPEAT 1941 1977 HEAT 14.		
FT	REPEAT 1982 2019 HEAT 15.		
FT	REPEAT 2020 2055 HEAT 16.		
FT	REPEAT 2057 2090 HEAT 17.		
FT	REPEAT 2319 2361 HEAT 18.		
FT	REPEAT 2379 2416 HEAT 19.		
SQ	SEQUENCE		2670 AA; 297333 MW; 948E9316D56D74C3 CRC64;
Query Match	Best Local Similarity 21.1%; Conservative 80%; Mismatches 190; Indels 141; Gaps 25;		
Matches	110;保守性 80%; 错配 190; 插入 141; 缺失 25;		
OY	10 VTKMLLAESSEVSVA-----ANAYMINSDMSDYLSAVDNFNE--RIGSQ 52		
Db	282 IEKALLRSPEIFITSGISSLALHGFADSKVDAASSLILSVLTSTVNGLASSNAEVRRNGQQ 341		
OY	53 VPKGSNCNSAVSAYMSRCAKQ-----DCITLQSLKYPEAKYQPL 92		
Db	342 TFKDSLANSADNESLSRVASSELITSLRTGKVTDQRYLFVDAULSSLSKHIDASMLINE 401		
OY	93 TLDPYQLEAFLKFKEEDDNP ANSTERFRWMFRRKQNSHFHDLYFNLLKNVTRDA 151		
RESULT	30		
Y066_NPVAC	STANDARD;	PRT;	808 AA.
ID	Y066_NPVAC		
AC	P41467;		
DT	01-NOV-1995 (Rel. 32, Created)		
DT	01-NOV-1995 (Rel. 32, Last sequence update)		
DT	01-NOV-1997 (Rel. 35, Last annotation update)		
DE	Hypothetical 94.0 kDa protein in poly-LEP3 intergenic region.		
OS	Autographa californica nuclear polyhedrosis virus (ACMV). Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Nucleopolyhedrovirus.		
OC	Nucleopolyhedrovirus.		
OX	NCBI_TAXID=46015;		
[1]			
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C6;		
RX	MEDLINE=93303173; PubMed=803024;		
RA	Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R. D.;		
RT	The complete DNA sequence of Autographa californica nuclear polyhedrosis virus.		
RL	Virology 202:586-605(1994).		
CC	-1-SIMILARITY: TO CORRESPONDING ORF IN OPMPV AND LDMPV.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> ) or send an email to license@isb-sib.ch).		
CC	EMBL: L22658; AAA66996.1; -.		
DR	EMBL: L22658; AAA66996.1; -.		
KW	Hypothetical protein.		
DOMAIN	106 117 POLY-B-PRO.		
SQ	SEQUENCE		808 AA; 93973 MW; 76A871D2B6633F8A CRC64;
Query Match	Best Local Similarity 4.5%; Score 105; DB 1; Length 808;		
Matches	99; 保守性 88%; 错配 182; 插入 118; 缺失 118; Gaps 25;		
OY	18 SESVSAANAYMINSDMSDYLSAVDNFNECOPKGSNCNSAVSAY--MSRCAKQ 74		
Db	368 SENVIKIQLERSNKLNDTIVQLRDAENRSEIQLQK----DLDEYKNMNLQNLN 422		
OY	75 CLTQSLKYPEAKYQPL--LTPPYQ---QLEAFILEK-SDANANSTEKRFW 124		











OY	254	-----VORLLATVAGYVDTPPW--YKKWY-----MKLK-----NF	280	730	744	ACTIN-BINDING ( POTENTIAL ).
Db	336	MRNVPPEKKTRGFLVATAGFENGWEGMDICWYGRHQNSSEGQQAADLKSQAAIDQING	395	277	280	POLY-GIX.
OY	281	MVNVRVITPKFFENKEIIRPSKALKRVSUTDKLLENKIGOGTVDF--FVKERIOPSKA	338	1471	1476	POLY-ALA.
Db	396	KLNRVIEKTNEKFHQIEKEFS---EVEGRIQD-ENVYEDTKIDLWNSYAEI-LVA	446	1482	1489	POLY-GLN.
OY	339	LKEKVSNDAKDLFENLIGQGVDFINNEIRPSKALRKVSTGAELFENKIGGVDFI	398	1493	1496	POLY-GLN.
Db	447	LE-----NOHTIDLTSEEMNLKFRTRROLRENAEDM---GSSCP---	483	1505	1511	POLY-GLN.
OY	399	NNEIIRDPSKALRKVTEADDLFENKIGOGTVDFI	431	119	119	METHYLATION ( TRT- ) ( POTENTIAL ).
Db	484	-----KIVHCDNACIESIRNGTYD	503	1577	1577	BAF2E8EBB#438C CRC64;
<b>RESULT 38</b>						
MISH_ACACA						
ID						
AC						
DT	01-FEB-1996 ( Rel. 33, Created)					
DT	01-OCT-2001 ( Rel. 40, Last sequence update)					
DE	High molecular weight form of myosin I ( HMWI ).					
OS	Acanthamoeba castellanii ( Ameba ).					
OC						
OX						
RN						
RP	SEQUENCE FROM N.A.					
RX	MEDLINE-91056121; PubMed-2243110;					
RA	Horowitz J.A., Hammer J.A. III;					
RT	"A new Acanthamoeba myosin heavy chain. Cloning of the gene and immunological identification of the polypeptide.";					
RL	J. Biol. Chem. 265:20646-20652(1990).					
CC	-!- SUBUNIT: MYOSIN I HEAVY CHAIN IS SINGLE-HEADED.					
CC	-!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.					
CC	-!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.					
CC	-!- SIMILARITY: CONTAINS 1 IO DOMAIN.					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).					
CC	-----					
DR	EMBL: M60954; AAA27709.1; -.					
DR	HSSP; P08799; IMD;					
DR	InterPro; IPR00048; IO.					
DR	InterPro; IPR00857; MYTH4.					
DR	InterPro; IPR01452; SH3.					
DR	InterPro; IPR001609; myosin_head.					
DR	Pfam: PF00612; IO; 1.					
DR	Pfam; PF00653; myosin_head; 2.					
DR	Pfam; PF00784; MYTH4; 1.					
DR	PRINTS; PR00193; MYOSINHEAVY.					
DR	PRODOM; PD000355; myosin_head; 1.					
DR	SMART; SM00015; IO; 1.					
DR	SMART; SM00242; MSC; 1.					
DR	SMART; SM00139; MYTH4; 1.					
DR	PROSITE; PS50002; SH3; 1.					
DR	PROSITE; PS50096; IO; 1.					
KW	Myosin: Actin-binding; ATP-binding; Methylation; Multigene family; SH3 domain.					
FT	DOMAIN 1 754 MYOSIN HEAD-LIKE.					
FT	DOMAIN 755 782 IO.					
FT	DOMAIN 1519 1577 SH3.					
FT	NP_BIND 168 175 ATP (POTENTIAL).					
FT	DOMAIN 628 650 ACTIN-BINDING (POTENTIAL).					
<b>RESULT 39</b>						
TOP2_SCHPO						
ID						
AC						
DT	01-AUG-1988 ( Rel. 08, Created)					
DT	16-OCT-2001 ( Rel. 40, Last sequence update)					
DE	DNA topoisomerase II ( EC 5.99.1.3 ).					
GN	TOP2 OR SPBC14.03C.					
OS	Eukaryote: Fungi: Ascomycota: Schizosaccharomycetes; Schizosaccharomycetales; Schizosaccharomyces; Schizosaccharomycetaceae; Schizosaccharomycetes.					
OC						
OC						
OC						
OX						
OX	NCBII_TaxID=4896;					
RN	SEQUENCE FROM N.A.					
RX	MEDLINE-87053875; PubMed=2023070;					
RA	Uemura T., Morikawa K., Yanagida M.;					
RT	"The nucleotide sequence of the fission yeast DNA topoisomerase II gene: structural and functional relationships to other DNA topoisomerases";					
RT	topoisomerase".					
RL	EMBO J. 5:2355-2361(1986).					
RN	[2]					



RX	MEDLINE-20085745;	PubMed-10619025;	FT	DOMAIN	524	719	LAMININ DOMAIN IV 1 (DOMAIN IV B).
RA	Hohenester E., Tiss D., Talts J.F., Timpl R.;	"The crystal structure of a laminin G-like module reveals the molecular basis of alpha-dystroglycan binding to laminins, perlecan, and agrin,"	FT	DOMAIN	720	752	LAMININ EGF-LIKE 5 (C-TERMINAL).
RT			FT	DOMAIN	753	802	LAMININ EGF-LIKE 6.
RT			FT	DOMAIN	803	860	LAMININ EGF-LIKE 7.
RT			FT	DOMAIN	861	913	LAMININ EGF-LIKE 8.
RL	Mol. Cell 4:783-792(1999)		FT	DOMAIN	914	952	LAMININ EGF-LIKE 9.
CC	-!- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION & ORGANIZATION OF CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.		FT	DOMAIN	963	1009	LAMININ EGF-LIKE 10.
CC	-!- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.		FT	DOMAIN	1010	1055	LAMININ EGF-LIKE 11.
CC	THE ALPHA-2 CHAIN IS A SUBUNIT OF LAMININ-2 (MEROSEN) AND LAMININ-4 (SMEROSIN).		FT	DOMAIN	1056	1101	LAMININ EGF-LIKE 12.
CC	-!- SUBCELLULAR LOCATION: EXTRACELLULAR.		FT	DOMAIN	1102	1161	LAMININ EGF-LIKE 13.
CC	-!- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR COMPONENT).		FT	DOMAIN	1162	1171	LAMININ EGF-LIKE 14 (N-TERMINAL).
CC	-!- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.		FT	DOMAIN	1172	1375	LAMININ DOMAIN IV 2 (DOMAIN IV A).
CC	-!- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.		FT	DOMAIN	1376	1415	LAMININ EGF-LIKE 14 (C-TERMINAL).
CC	-!- DISEASE: DEFECTS IN LAMA2 ARE A CAUSE OF MURINE MUSCULAR DYSTROPHY (DY2J).		FT	DOMAIN	1416	1464	LAMININ EGF-LIKE 15.
CC	-!- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).		FT	DOMAIN	1465	1522	LAMININ EGF-LIKE 16.
CC	-!- SIMILARITY: CONTAINS 17 LAMININ EGF-LIKE DOMAINS.		FT	DOMAIN	1523	1569	LAMININ EGF-LIKE 17.
CC	-!- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.		FT	DOMAIN	1570	2140	DOMAIN II AND I.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).		FT	DOMAIN	2141	2324	LAMININ DOMAIN IV 1.
CC			FT	DISULFID	2336	2517	LAMININ EGF-LIKE 2.
CC			FT	DISULFID	2522	2706	LAMININ G-LIKE 3.
CC			FT	DISULFID	2759	2930	LAMININ G-LIKE 4.
CC			FT	DISULFID	2929	3106	LAMININ G-LIKE 5.
CC			FT	DISULFID	1662	1863	COILED COIL (POTENTIAL).
CC			FT	DISULFID	1923	2145	COILED COIL (POTENTIAL).
DR	EMBL; U12147; AAC52165.1.		FT	DISULFID	283	292	LAMININ G-LIKE 1.
DR	EMBL; X69869; CAA49302.1.		FT	DISULFID	285	303	LAMININ G-LIKE 2.
DR	EMBL; S75315; AAB33573.1.		FT	DISULFID	305	314	LAMININ G-LIKE 3.
DR	PDB; 1Q0U; 03-DEC-99.		FT	DISULFID	317	337	LAMININ G-LIKE 4.
DR	MGD; MGIR:9912; Lama2.		FT	DISULFID	340	349	LAMININ G-LIKE 5.
DR	Inter-Pro: IPR00561; EGF-Like.		FT	DISULFID	342	374	COILED COIL (POTENTIAL).
DR	InterPro; IPR01886; Laminin.		FT	DISULFID	377	385	LAMININ G-LIKE 1.
DR	InterPro; IPR00034; Laminin_B.		FT	DISULFID	389	407	LAMININ G-LIKE 2.
DR	Inter-Pro; IPR002049; Laminin_EGF.		FT	DISULFID	410	422	LAMININ G-LIKE 3.
DR	InterPro; IPR001791; Laminin_G.		FT	DISULFID	412	438	LAMININ G-LIKE 4.
DR	Inter-Pro; IPR00053; Laminin_EGF.		FT	DISULFID	440	449	LAMININ G-LIKE 5.
DR	pfam; PF00054; Laminin_G.		FT	DISULFID	452	462	COILED COIL (POTENTIAL).
DR	pfam; PF00055; Laminin_Nterm.		FT	DISULFID	465	478	LAMININ G-LIKE 1.
DR	PRINTS; PRO0011; EGFLAMININ.		FT	DISULFID	467	482	LAMININ G-LIKE 2.
DR	PRODOM; PPD002082; Laminin.		FT	DISULFID	484	493	LAMININ G-LIKE 3.
DR	PRODOM; PPD00331; Laminin_B.		FT	DISULFID	496	511	LAMININ G-LIKE 4.
DR	SMART; SM00180; EGF_Lam.		FT	DISULFID	753	762	LAMININ G-LIKE 5.
DR	SMART; SM0001; EGF_Like.		FT	DISULFID	755	769	COILED COIL (POTENTIAL).
DR	SMART; SM00281; LamB.		FT	DISULFID	772	781	LAMININ G-LIKE 1.
DR	SMART; SM00282; LamG.		FT	DISULFID	784	800	LAMININ G-LIKE 2.
DR	SMART; SM00136; LamNt.		FT	DISULFID	803	818	LAMININ G-LIKE 3.
DR	PROSITE; PS00022; EGF_1.		FT	DISULFID	805	828	LAMININ G-LIKE 4.
DR	PROSITE; PS00186; EGF_2.		FT	DISULFID	831	840	LAMININ G-LIKE 5.
DR	PROSITE; PS001248; LAMININ_TYPE-EGF.		FT	DISULFID	843	858	COILED COIL (POTENTIAL).
DR	PROSITE; PS50025; LAMIN_DOMAIN.		FT	DISULFID	861	875	LAMININ G-LIKE 1.
KW	Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil; Laminin EGF-like domain; Cell adhesion; Repeat; Signal; 3D-structure.		FT	DISULFID	883	882	LAMININ G-LIKE 2.
KW	SIGNAL		FT	DISULFID	897	911	LAMININ G-LIKE 3.
FT	1		FT	DISULFID	914	926	LAMININ G-LIKE 4.
FT	22		FT	DISULFID	916	933	LAMININ G-LIKE 5.
FT	LAMININ ALPHA-2 CHAIN		FT	DISULFID	935	944	COILED COIL (POTENTIAL).
FT	LAMININ N-TERMINAL (DOMAIN VI).		FT	DISULFID	947	960	LAMININ G-LIKE 1.
FT	LAMININ EGF-LIKE 1.		FT	DISULFID	963	975	LAMININ G-LIKE 2.
FT	LAMININ EGF-LIKE 2.		FT	DISULFID	965	981	LAMININ G-LIKE 3.
FT	LAMININ EGF-LIKE 3.		FT	DISULFID	983	992	LAMININ G-LIKE 4.
FT	LAMININ EGF-LIKE 4.		FT	DISULFID	995	1007	LAMININ G-LIKE 5.
FT	PS00186; EGF_2.		FT	DISULFID	1010	1019	COILED COIL (POTENTIAL).
FT	PS001248; LAMININ_TYPE-EGF.		FT	DISULFID	1012	1026	LAMININ G-LIKE 1.
FT	PS50025; LAMIN_DOMAIN.		FT	DISULFID	1028	1037	LAMININ G-LIKE 2.
FT	DISULFID		FT	DISULFID	1040	1053	LAMININ G-LIKE 3.
FT	DISULFID		FT	DISULFID	1056	1068	LAMININ G-LIKE 4.
FT	DISULFID		FT	DISULFID	1058	1075	LAMININ G-LIKE 5.
FT	DISULFID		FT	DISULFID	1077	1085	COILED COIL (POTENTIAL).
FT	DISULFID		FT	DISULFID	1089	1099	LAMININ G-LIKE 1.
FT	DISULFID		FT	DISULFID	1416	1425	LAMININ G-LIKE 2.
FT	DISULFID		FT	DISULFID	1418	1432	LAMININ G-LIKE 3.
FT	DISULFID		FT	DISULFID	1435	1444	LAMININ G-LIKE 4.
FT	DISULFID		FT	DISULFID	1447	1462	LAMININ G-LIKE 5.
FT	DISULFID		FT	DISULFID	1465	1480	COILED COIL (POTENTIAL).

Query Match 4.3%; Score 102.5; DB 1; Length 3106;  
 Best Local Similarity 20.3%; Pred. No. 85;  
 Matches 109; Conservative 68; Mismatches 178; Indels 183; Gaps 26;

Qy 4 SDSVGDVYKTLASESESVDAAVYMINSDMSDYLAVSAVDNAERICSQVPGNSCASY 63  
 Db 1671 AEELEEFTRGLVDDAEINE--KAVKLNELGN---ODKTAERNELEQLK---EI 1717  
 Qy 64 SAYMRSRCAKQDCCLTQLSKYPLEAKYQPLTPDPYQLEAFLFKESDANPANSTEKFW 123  
 Db 1718 DRMKLKELRSKDQTL-----OKEVA---EDELVAAEGLLKRVN 1751  
 Qy 124 MRF--RRGKNHSYFHDLYVNLLR-KNVTRDA---DATDIEFAASRYLYMATLYKYT 175  
 Db 1752 KLGEGPRAQNDMEKLQDQKLARYNKNDLQADMLREATDKTRDANR--LSAANQKNT 1808  
 Qy 176 NVBFGASPFENKLSTFTGLFGWSIKRALQQT-SNLPDIGTEHSVSRLQHTTSSYKD 234  
 Db 1809 ILEKKKEAT-----EGSKRQENTLKGENDID---EANQLGEINSVIDY 1851  
 Qy 235 MDTQIPALPKFAKFRSLV-----VORLL-----ATVAGYYVDPW 269  
 Db 1852 VDDITKTLPPMSEBELSDKIDDIAQEIKDRRLAEWFOAESHAQQLNDSSAVLDGILD-- 1908  
 Qy 270 YKKWYMKLKNFMN-----RVFPPTKKPNKERREPSKALKEWVSTDVKDFENKIGCT 324  
 Db 1909 -----EARNSFNATAAFRAYSNIKYIDE-----AEKVARAEKEL---AQGA 1948  
 Qy 325 VDFFENKEIRDPSKALKE-----KVSNDAK---DLFENKISGGTVDFI 363  
 Db 1949 -----TKLATSPQGLLKEDKGSKQKSFRILNEAKVNGHNDLNLKTRETADR 2004  
 Qy 364 N-----NEIRDESKALIRKVSTGAEDLFENKIGQGTVDFINNEIRDPSKALIRKVYTE 416  
 Db 2005 NSGLIGALNDTMDKLSATNDTAAKQAYREKA-----FREANDTAKAVLAQV-- 2051  
 Qy 417 ADDLFENKIG-----OGTVPPEINKEIRDPSK-----ALIRKYSTEAUNLLEK 458  
 Db 2052 -KDLHQNLQDGKQNYKLADSVARTNAVYKDPSKNKIADAGTSVRNLEQEAIDLIDK 2108

Search completed: October 19, 2002, 07:49:52  
 Job time : 36 secs